UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 09/934,799 | 08/23/2001 | Jason Peter Andrew Charlesworth | 1263.1203 | 9079 |

5514 7590 08/14/2003

FITZPATRICK CELLA HARPER & SCINTO
30 ROCKEFELLER PLAZA
NEW YORK, NY 10112

| EXAMINER | |
|---|---|
| STORM, DONALD L | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2654 | 14 |

DATE MAILED: 08/14/2003

Please find below and/or attached an Office communication concerning this application or proceeding.

| Office Action Summary | Application No. 09/934,799 | Applicant(s) CHARLESWORTH ET AL. 2 |
|---|---|---|
| | Examiner Donald L. Storm | Art Unit 2654 |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *28 January 2003 and 12 June 2003* .

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *98-121* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *98-121* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *23 August 2001* is/are: a)☐ accepted or b)☒ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

**Priority under 35 U.S.C. §§ 119 and 120**

13)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☒ All b)☐ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449) Paper No(s) *11* .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

## DETAILED ACTION

1. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

### *Priority*

2. Papers submitted June 12, 2003 (paper 12) under 35 U.S.C. 119(a)-(d) have been received.

### *Information Disclosure Statement*

3. A copy of the search report of the European Patent Office for Application No. 01304243.7 and the copies of the documents are present in the application file, and they have been considered by the Examiner.

### *Drawings*

4. The drawings are objected to under 37 CFR § 1.83(a) because they fail to clearly show significant features of the subject matter sought to be patented. See MPEP § 608.02(d). At a minimum, representation of the following features should be added to the drawings to show the claimed invention as a whole:

a. annotation storage areas comprising block storage areas (claims 98, 117),

b. block storage areas having a time index identifying the block's timing within the annotation (claims 98, 117),

c. block storage areas including node storage areas (claims 98, 117),

d. node storage areas associated with nodes representing beginning or ending points in time within the annotation (claims 98, 117),

e. node storage areas having time offset storage areas (claims 98, 117),

f. node storage areas having phoneme links storage areas and word link storage areas (claims 98, 108, 109, 110, 117),

g. phoneme link storage areas having phoneme storage areas (claims 98, 117),

h. word link storage areas having word storage areas (claims 98, 117);

i. a phoneme link storage area having a node offset storage area (claim 102);

j. a word link storage area having a node offset storage area (claim 103);

k. phoneme storage areas having a weighting storage area (claim 104);

l. word link storage areas having a weighting storage area (claim 106); and

m. annotation storage areas comprising a header storage area (claim 115);

n. header storage identifying ASR language, phoneme, and word sets (claim 116); and

o. a data structure arranged in a time ordered sequence of blocks (claim 118, and others).

***Specification***

5. The title is objected to because it is not sufficiently descriptive of the invention. A new title is required that is clearly indicative of the invention to which the claims are directed. See MPEP § 606.01. The Examiner suggests that the Applicant consider a title including these elements: "Annotation Database Block Areas to Accommodate Time-Indexed Nodes, Phoneme Links, and Word Links."

6. The specification is objected to using the same rationale as in the prior Office action (paper 10) as failing to provide proper antecedent basis for the following claimed subject matter: data arranged in blocks of equal time duration (at least claim 99). Appropriate correction is required.

***Claim Informalities***

7. Claim 100 is objected to under 37 CFR 1.75(a) because the meaning of the phrase "the time index associated with each block storage area" needs clarification. Because no time index associated with a block storage area was previously recited, it is unclear as to what element this phrase is making reference. The time index identified a block in claim 98. To further timely prosecution and evaluate prior art, the Examiner has interpreted this phase to refer to --the time index associated with each temporal block--.

8. Claim 100 is objected to under 37 CFR 1.75(a) because the meaning of the phrase "the database" needs clarification. Because no database was previously recited, it is unclear as to what element this phrase is making reference. To further timely prosecution and evaluate prior art, the Examiner has interpreted this phase to refer to --the data file--.

9. Claim 113, and by dependency claim 114, are objected to under 37 CFR 1.75(a) because the meaning of the phrase "said audio and/or video signal" needs clarification. Because no audio and/or video signal was previously said, it is unclear as to what element this phrase is making reference. It might be somehow referring indirectly to the time sequential signal of claim 98. To further timely prosecution and evaluate prior art, the Examiner has interpreted this phase to refer to --said audio and/or video data-- of claim 112.

10. Claim 114 is objected to under 37 CFR 1.75(a) because the meaning of the phrase "said annotation" needs clarification. Because there is annotation representing the time sequential signal of claim 98 and annotation associated with said audio and/or video signal and derived from said audio and/or video data of claim 113, it is unclear as to what element this phrase is making reference. To further timely prosecution and evaluate prior art, the Examiner has interpreted this phase to refer to --the annotation derived from said audio and/or video data--.

11. The Examiner notes, without objection, the possibility of informalities in the claims. The Applicant may wish to consider changes during normal review and revision of the disclosure.

a. In claim 107, line 3, should the word "is" be --it--?

b. In claim 115, line 3, if the plural noun "areas" is intended to be the subject of the singular verb "comprises", it does not agree in number.

c. In claim 117, lines 1-2, is a word or phrase missing from the phrase "computer readable medium storage computer executable instructions for defining a data structure of ruse in accessing"?

12. The Examiner notes, without objection, that the following phrases assume inherent antecedence to previously recited claim elements: (claim 98) "the corresponding annotation"; (claim 98), "the corresponding block"; (claim 98) "the corresponding node"; (claim 101) "the corresponding data file"; (claim 102) "the phoneme data"; (claim 102) "the corresponding phoneme"; (claim 103) "the corresponding word"; (claim 104) "the corresponding phoneme"; (claim 106) "the corresponding word"; (claim 113) "the annotation associated with

said. . . .signal"; (claim 115) "said annotations"; (claim 115) "the annotation"; (claim 117) "the corresponding annotation"; (claim 117), "the corresponding block"; (claim 117); "the corresponding node". The following phrases provide an inherent reference by their recitation: (claim 116) "the language and the word and phoneme sets". The Applicant may wish to consider if the phrases recite the claimed subject matter that the Applicant wants.

### *Claim Rejections - 35 USC § 101*

13. Claims 98-116 and 118-121 are rejected under 35 U.S.C. 101 because the claimed invention is directed to non-statutory subject matter.

14. Regarding claim 98, data files, annotations, phoneme data, word data, indexes to data, time information, associations among data, and nodes comprise only nonfunctional descriptive material, even when stored on a recording medium. Data *per se* comprises only nonfunctional descriptive material. It is not structurally and functionally interrelated to the medium, but merely carried by the medium. Nonfunctional descriptive material includes the compilation or mere arrangement of data. A data structure *per se*, data storage areas adapted to data *per se*, storage areas containing data *per se*, storage areas comprising other storage areas, and storage areas' relative positions comprise only nonfunctional descriptive arrangements of the data compilation. Purely nonfunctional descriptive material is nonstatutory despite the fact that it might inherently have some usefulness. Such descriptive material is not statutory because it is neither a useful process, machine, manufacture, nor composition of matter. Purely nonfunctional, descriptive material cannot alone provide a practical application. Data fields and descriptive data that they contain, no matter what name or arrangement is given to them, cannot alone change a general

purpose computer into a special purpose machine by activating electrical paths and deactivating other paths. All claim limitations have been considered, and the claimed data has been found nonstatutory as a mere arrangement of data.

The further limitations of the dependent claims 99-116 continue to describe data and the arrangement of data; they do not provide any statutory category of invention having a practical application to satisfy the requirements of 35 U.S.C. 101.

15. Regarding claim 118, data, nodes, and links of a lattice comprise only nonfunctional descriptive material. Data *per se* comprises only nonfunctional descriptive material. Data associating phonemes and at least one word with links of a lattice merely provide an arrangement of the data compilation. Nonfunctional descriptive material includes the compilation or mere arrangement of data. A data structure *per se* and an arrangement of positions of component blocks of a data structure comprise only nonfunctional descriptive arrangements of data. Purely nonfunctional descriptive material is nonstatutory despite the fact that it might inherently have some usefulness. Such descriptive material is not statutory because it is neither a useful process, machine, manufacture, nor composition of matter. Purely nonfunctional descriptive material cannot alone provide a practical application. All claim limitations have been considered, and the claimed data has been found nonstatutory as a mere arrangement of data.

The further limitations of the dependent claims 119-121 continue to describe data and the arrangement of data; they do not provide any statutory category of invention having a practical application to satisfy the requirements of 35 U.S.C. 101.

## *Claim Rejections - 35 USC § 103*

16. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

***Bird et al.***

17. Claims 98-103, 108-113, and 117-120 are rejected under 35 U.S.C. 103(a) as being unpatentable over Steven Bird and Mark Liberman, "Towards a Formal Framework for Linguistic Annotations," 1998 (Bird et al.), already of record.

18. Regarding claim 98, Bird et al. [at section 3] makes obvious the claimed limitations as a whole recognizable to one versed in the art because the reference explicitly describes the following elements and suggests their combination into a common core of annotations:

a data structure comprising annotation storage areas [at section 7, as a file encoding an annotated graph];

annotation storage areas comprising block storage areas [at section 3.1, as files containing sets of lines 2360 through 12257, etc. and sets of lines 2360 through 49066, etc.];

block storage areas having a time index identifying the block's timing within the annotation [at section 3.2, as 2360 and 49066 information about the time and, for example, 12257 time between them];

block storage areas including node storage areas [at section 3.1, as 2360 through 5200 times are attributes of nodes];

node storage areas associated with nodes representing beginning or ending points in time of a word and/or a phoneme within the annotation [at section 3.1, as 2360 and 5200 times are starting and ending times of word "she"];

node storage areas having time storage areas [at section 3.1, as lines showing times 0, 2360, 3720, 5200, etc.];

the time storage areas are time offset areas [at section 4.8, as time is offset];

node storage areas having phoneme link storage areas and word link storage areas [at section 3.1, as lines showing phonetic transcription h#, sh, iy, etc. and lines showing words she, had, your, etc. and the phoneme or word label is a property of an edge connecting nodes];

phoneme link storage areas having phoneme storage areas [at section 3.1, as phonetic transcription h#, sh, iy, etc.];

word link storage areas having word storage areas [at section 3.1, as words she, had, your, etc.];

the annotation storage areas correspond with data files for accessing [at section 8, as annotation encoded in files for search];

the annotation storage areas contain an annotation representing a time sequential signal [at section 3.1, as the file contains 2360 through 49066 information about the time combined with a string of words];

block storage areas contain phoneme and word data forming a temporal block of the annotation [at section 3.1, as phonetic transcription h#, sh, iy, etc. and words she, had, your, etc. containing 2360 through 12257 and 49066 information about the time combined with a string of words];

time offset areas contain a time offset defining a time of the node relative to the time index of its block [at section 5.2, as annotation vertices temporally constrained to an assigned temporal locus of a semi-anchored vertex with reference to section 3];

phoneme storage areas containing data identifying a phoneme associated with the node or word storage areas containing data identifying a word associated with the node [at section 3.1, as phonetic transcription h#, sh, iy, etc. for the respective time information 0, 2360, 3720, etc.]; and

one or more of the node storage areas has a phoneme link storage area and a word link storage area [at section 3.1 as line 2360 through 5200 has phonetic transcriptions "sh", "iy", and has word "she"].

To the extent that Bird et al. does not explicitly describe one common arrangement displaying all of the data characteristics and particular arrangements recited in the claim, Bird et al. suggests that it may de done. Bird et al. [at sections 7 and 8] recognizes the need to customize the described conceptual framework to satisfy particular applications. Because Bird et al. invites the user to make changes and combinations, it would have been obvious to one of ordinary skill in the art of data storage to selectively choose between configurations consisting of Bird et al.'s known elements and achieve a desired combination of advantages by selecting the appropriate ones of Bird et al.'s descriptions to support specific applications, for example search.

19. Regarding claim 99, Bird et al. also describes:

temporal blocks of equal time duration [at section 3.1 and Fig. 1, as, for example, the block 3/5200 -> 4/6100 -> 5/8720 -> 6/9680 equals the block 3/5200 -> 6/9680 in duration].

20. Regarding claim 100, Bird et al. also describes:

the time index identifies the block storage area location within the database *{Exmr: the data file}* [at section 6, as indices specify where to find annotation graphs of sets of nodes having the same time reference].

21. Regarding claim 101, Bird et al. also describes:

a data file representative of a time sequential signal [ at section 4.8, as the time-function file];

the time offsets are time synchronised with the time sequential signal [at section 4.8, as resolution of time references are mediated to reference the time function file].

22. Regarding claim 102, Bird et al. also describes:

the phoneme data identifies a phoneme which starts at a node and ends at another node [at section 3.1, as phonetic transcriptions "sh" starts at 2360 and ends at 3720];

the phoneme link storage area has a node storage area containing data identifying the node at which the phoneme ends [at section 3.1, as lines showing phonetic transcription h#, sh, iy, etc. have node storage areas 2360, 3720, 5200 etc. containing 2360, 3720, 5200, etc. information about the time];

the node storage area is offset storage [at section 4.8, as time is offset].

23. Regarding claim 103, Bird et al. also describes:

the word data identifies a word which begins at a node and ends at another node [at section 3.1, as words she, etc. begins at 2360 and ends at 5200, etc];

the word link storage area has a node storage area containing data identifying the node at which the word ends [at section 3.1, as lines showing words she, etc. have node storage areas 5200, etc. containing 5200, etc. information about the time];

the node storage area is offset storage [at section 4.8, as time is offset].

24. Regarding claim 108, Bird et al. also describes:

a node storage area [at section 3.1, as 2360 line showing phonetic transcription, 3720 line showing phonetic transcription, etc.];

a phoneme link storage area [at section 3.1, as phonetic transcription sh, phonetic transcription iy, etc.];

a node storage area includes a plurality of phoneme link storage areas [at section 3.6, as a single line of multiple lines collapsed for compactness].

25. Regarding claim 109, Bird et al. also describes:

a node storage area [at section 3.1, as 2360 line showing word, 5200 line showing word, etc.];

a word link storage area [at section 3.1, as word "she", word "had", etc.];

a node storage area includes a plurality of word link storage areas [at section 3.6, as a single line of multiple lines collapsed for compactness].

26. Regarding claim 110, Bird et al. also describes:

a node storage area [at section 3.1, as 2360 line showing word, 5200 line showing word, etc.];

a phoneme link storage area [at section 3.1, as phonetic transcription sh, phonetic transcription iy, etc.];

a word link storage area [at section 3.1, as word "she", word "had", etc.];

a node storage area includes a plurality of word link storage areas [at section 3.6, as a single line of multiple lines collapsed for compactness, for example line showing 5 0 (phonetic D) 1 (phoneme D) 2 (word "the"), etc, ].

27. Regarding claim 111, Bird et al. also describes:

a plurality of data file storage areas each storing a data file [at section 4.8, as a set of time-function files of a corpus].

28. Regarding claim 112, Bird et al. also describes:

a data file of audio and/or video data [at section 1, as recorded signals of audio, video].

29. Regarding claim 113, Bird et al. also describes:

annotation associated with audio and/or video [at section 1, as annotations describing audio, video];

and derived from the audio and/or video data [at section 1, as annotation is an orthographic transcription of speech].

30. Claim 117 sets forth limitations similar to claim 98. Bird et al. describes the limitations as indicated there. Bird et al. also describes;

computer instructions for defining the data structure [at section 1, as software for the creation of the database].

However, Bird et al. does not describe any apparatus to store and execute the software. In particular, Bird et al. does not explicitly describe computer readable medium storage computer executable instructions for defining the structure.

To the extent that a computer is not necessarily in Bird et al.'s system so that the software may be stored and executed, it would have been obvious to one of ordinary skill in the art of implementing functional descriptions of operations at the time of invention to include computer readable media storing executable program instructions to implement the processing functions of Bird et al. in software because computer implementation of executable software modules would eliminate the tedium of manual calculation of repetitive operations.

31. Regarding claim 118, Bird et al. [at section 4.5] makes obvious the claimed limitations as a whole recognizable to one versed in the art because the reference explicitly describes the following elements:

data defining nodes within a lattice [at section 3.1, as lines of numbers and labels in the corpus interpreting times as nodes in a directed, acyclic graph];

data defining links connecting the nodes within a lattice [at section 3.1, as lines of labels and numbers in the corpus interpreting labels as edges which connect nodes in a directed, acyclic graph];

data associating phonemes with the links [at section 3.1, as type of the edge label P for phonetic transcription]; and

data associating words with links [at section 3.1, as type of the edge label W for word];

data structure is arranged in a time ordered sequence of blocks [at section 3.5, as the structure of the annotation is divided into sections containing segments at times in order 4.233, 8.015, 11.040, etc.].

To the extent that Bird et al. does not explicitly describe one common arrangement displaying all of the data characteristics and particular arrangements recited in the claim. However, Bird et al. [at sections 7 and 8] recognizes the need to customize conceptual framework to satisfy particular applications. Because Bird et al. invites the user to make changes and combinations, it would have been obvious to one of ordinary skill in the art of data storage to selectively choose between configurations consisting of Bird et al.'s known elements and achieve a desired combination of advantages by selecting the appropriate ones of Bird et al.'s descriptions to support specific applications, for example search.

32. Regarding claim 119, Bird et al. also describes:

each node comprises time offset of a node from the start point of a block [at section 5.2, as annotation vertices temporally constrained to an assigned temporal locus of a semi-anchored vertex with reference to section 3];

33. Regarding claim 120, Bird et al. also describes:

a header having start time [at section 3.5, as <Segment S_time=4.233].

***Bird et al. and James et al.***

34. Claims 104-107 and 114 are rejected under 35 U.S.C. 103(a) as being unpatentable over Steven Bird and Mark Liberman, "Towards a Formal Framework for Linguistic Annotations," 1998 (Bird et al.), in view of D. A. James and S. J. Young, "A Fast Lattice-Based Approach to Vocabulary Independent Wordspotting," 1994 (James et al.), both already of record.

35. Regarding claim 104, Bird et al. describes the phoneme storage areas as for claim 98.

Bird et al. [at section 4.6] describes that the annotation framework will allow the arcs to carry multiple attributes and values. Bird et al. does not explicitly provide weights for the phonemes and words on the edges of the lattice that provide a measure of confidence.

James et al. [at section 2] also describes a lattice framework and designs it to provide reliable search results quickly. So that the search result also returns a measure of its reliability, James et al. describes:

a weighting for the phoneme [at section 2, as a likelihood score for the phone labeling the edge].

Since both Bird et al. and James et al. annotate databases with similar lattices applied to searching the database, it would have been obvious to one of ordinary skill in the art of search and retrieval at the time of invention to add James et al.'s concept of likelihood for the edge labels to be able to provide weights when Bird et al.'s edge labels are matched during search because that would allow the user or further applications to assess how closely the results of the search compare to the search query.

James et al. [at section 2] also stores the phone hypothesis for manipulation by computer. However, neither Bird et al. nor James et al. explicitly describes storage areas for the weighting.

The many teachings throughout Bird et al. devoted to storing the values represented in the phoneme and word lattice make it obvious to one of ordinary skill in the art of storing particular values corresponding to parameters to provide the phoneme storage area with a weighting storage area suitable for storing James et al.'s phoneme weight values in the same manner that Bird et al. provides storage areas for other values that are explicitly stored because storage of the weights with corresponding phonemes would allow the user or further applications to assess how closely the results of James et al.'s speech recognition identify the input speech.

36. Regarding claim 105, James et al. also describes:

the phoneme and word data is generated by an ASR system [at section 2, as phone lattices are generated by a speech recognizer and word spotting is performed to obtain a path for a keyword through the lattice];

the weighting represents a confidence output of correct recognition of the phoneme by the ASR [at section 2, as a likelihood score for the phone hypothesis by the speech recognizer].

37. Regarding claim 106, Bird et al. describes the word link storage areas as for claim 98.

Bird et al. [at section 4.6] describes that the annotation framework will allow the arcs to carry multiple attributes and values. Bird et al. does not explicitly provide weights for the phonemes and words on the edges of the lattice that provide a measure of confidence.

James et al. [at section 2] also describes a lattice framework and designs it to provide reliable search results quickly. So that the search result also returns a measure of its reliability, James et al. describes:

a weighting for the word [at section 2, as the cumulative DP matching score for keyword phones].

Since both Bird et al. and James et al. annotate databases with similar lattices applied to searching the database, it would have been obvious to one of ordinary skill in the art of search and retrieval at the time of invention to add James et al.'s concept of likelihood for the edge labels to be able to provide weights when Bird et al.'s word edge labels are matched during word searches because that would allow the user or further applications to assess how closely the results of the search compare to the search query.

James et al. [at section 2] also stores the phone hypothesis for manipulation by computer. However, neither Bird et al. nor James et al. explicitly describes storage areas for the weighting.

The many teachings throughout Bird et al. devoted to storing the values represented in the phoneme and word lattice make it obvious to one of ordinary skill in the art of storing particular values corresponding to parameters to provide the word storage area with a weighting storage area suitable for storing James et al.'s keyword cumulative weight values of phoneme paths in the same manner that Bird et al. provides storage areas for other values that are explicitly stored because storage of the weights with corresponding words would allow the user or further applications to assess how closely the results of James et al.'s speech recognition identify the input speech.

38. Regarding claim 107, James et al. also describes:

the phoneme and word data is generated by an ASR system [at section 2, as phone lattices are generated by a speech recognizer and word spotting is performed to obtain a path for a keyword through the lattice];

the weighting represents a confidence output of correct recognition of the word by the ASR [at section 2, as a cumulative DP matching score of phone hypotheses by the speech recognizer for the keyword path].

39. Regarding claim 114, Bird et al. also describes:

annotation derived from the audio and/or video data [at section 1, as annotation is an orthographic transcription of speech].

Bird et al. [at section 1] also describes software for the creation of the database. However, Bird et al. does not further describe the software that should be used. In particular, Bird et al. does not explicitly describe automatic speech recognition generating phoneme and words. Moshier [at title] describes speech recognition, but does not explicitly describe annotations for a database of speech.

James et al. [at section 2] also describes a lattice framework and designs it to provide reliable search results quickly. So that the search result also returns a measure of its reliability, James et al. describes:

the annotation is derived from audio and/or video [at section 1, as automatically indexing video and unknown speech];

the annotation is derived by an ASR system [at section 2, as phone lattices are generated by a speech recognizer and word spotting is performed to obtain a path for a keyword through the lattice].

Since both Bird et al. and James et al. annotate databases with similar lattices applied to searching the database, an artisan seeking to implement Bird et al.'s automatic generation of phonemes and words would have found it obvious at the time of invention to use James et al.'s

concept of speech recognition to convert the audio signal comprising speech to the phonemes and words that can populate Bird et al.'s lattice because automatic processor implementation of the algorithm would eliminate the tedium of manual calculation of repetitive operations.

***Bird et al. and James et al. and Moshier***

40. Claims 115 and 121 are rejected under 35 U.S.C. 103(a) as being unpatentable over Steven Bird and Mark Liberman, "Towards a Formal Framework for Linguistic Annotations," 1998 (Bird et al.) in view of D. A. James and S. J. Young, "A Fast Lattice-Based Approach to Vocabulary Independent Wordspotting," 1994 (James et al.), both already of record, and in view of Moshier [US Patent 4,227,176].

41. Regarding claim 115, Bird et al. also describes:

the speech data as audio [at section 1, as speech included as a kind of audio];

header storage area containing header data [at section 3.4, as <header> with title and sound file reference].

Bird et al. [at section 1] also describes software for the creation of the database. However, Bird et al. does not further describe the software that should be used. In particular, Bird et al. does not explicitly describe automatic speech recognition generating phoneme and words. Moshier [at title] describes speech recognition, but does not explicitly describe annotations for a database of speech.

James et al. [at section 2] also describes a lattice framework and designs it to provide reliable search results quickly. So that the search result also returns a measure of its reliability, James et al. describes:

the annotation is derived from an ASR system [at section 2, as phone lattices are generated by a speech recognizer and word spotting is performed to obtain a path for a keyword through the lattice].

Since both Bird et al. and James et al. annotate databases with similar lattices applied to searching the database, an artisan seeking to implement Bird et al.'s automatic generation of phonemes and words would have found it obvious at the time of invention to use James et al.'s concept of speech recognition to convert the audio signal comprising speech to the phonemes and words that can populate Bird et al.'s lattice because automatic processor implementation of the algorithm would eliminate the tedium of manual calculation of repetitive operations.

Both Bird et al. and James et al. store the phone hypothesis for manipulation by computer. However, neither Bird et al. nor James et al. explicitly describes storage areas in the header for data relating to the ASR system.

Moshier [at column 3, lines 38-56] also describes speech recognition of words using a sequence of phonemes. Like Bird et al. and James et al., Moshier [at columns 22-26] describes storage areas for the recognition results and describes likelihood values to measure correctness of recognition. Moshier also describes:

a header storage area containing header data relating to the ASR system [at column 23, lines 5-14 and diagram, as saving the time and likelihood of a detected peak for the current word pattern sequence].

In view of Bird et al.'s express desire to keep the word and phoneme lattice conceptually simple, it would have been obvious to one of ordinary skill in the art of data storage at the time of invention to adopt Moshier's concept of storing the ASR likelihood of recognition in a header area because that would preserve the likelihood values that would allow the user or further applications

to assess how closely the results of the search compare to the search query, but would keep the phoneme and word storage areas compact so that further application processing, such as search, would not be slowed.

42. Regarding claim 121, Bird et al. also describes:

header information [at section 3.4, as <header> with title and sound file reference];

information of word [at section 3.1, as words she, etc.].

Bird et al. [at section 1] also describes software for the creation of the database. However, Bird et al. does not further describe the software that should be used. In particular, Bird et al. does not explicitly describe automatic speech recognition generating phoneme and words. Moshier [at title] describes speech recognition, but does not explicitly describe annotations for a database of speech.

James et al. [at section 2] also describes a lattice framework and designs it to provide reliable search results quickly. So that the search result also returns a measure of its reliability, James et al. describes:

information of word [at section 2, as word spotting obtains a path for a keyword through the lattice].

Since both Bird et al. and James et al. annotate databases with similar lattices applied to searching the database, an artisan seeking to implement Bird et al.'s automatic generation of phonemes and words would have found it obvious at the time of invention to use James et al.'s concept of speech recognition to convert the audio signal comprising speech to the phonemes and words that can populate Bird et al.'s lattice because automatic processor implementation of the algorithm would eliminate the tedium of manual calculation of repetitive operations.

Both Bird et al. and James et al. store the phone hypothesis for manipulation by computer. However, neither Bird et al. nor James et al. explicitly describes storage areas in the header for data relating to the ASR system.

Moshier [at column 3, lines 38-56] also describes speech recognition of words using a sequence of phonemes. Like Bird et al. and James et al., Moshier [at columns 22-26] describes storage areas for the recognition results and describes likelihood values to measure correctness of recognition. Moshier also describes:

a header storage comprising information of word (or of something else) [at column 23, lines 5-14 and diagram, as saving the time and likelihood of a detected peak for the current word pattern sequence].

In view of Bird et al.'s express desire to keep the word and phoneme lattice conceptually simple, it would have been obvious to one of ordinary skill in the art of data storage at the time of invention to adopt Moshier's concept of storing the ASR likelihood of recognition in a header area because that would preserve the likelihood values that would allow the user or further applications to assess how closely the results of the search compare to the search query, but would keep the phoneme and word storage areas compact so that further application processing, such as search, would not be slowed.

***Allowable Subject Matter***

43. Claim 116 recites subject matter that is not taught or made obvious by prior art when considered with the limitations of the base claim and intervening claims.

Claim 116 requires that a header storage area identify the language, the word set, and the phoneme set, all of which were used by the ASR system, and that ASR system generated the

annotation. The closest prior art (James et al., Moshier) describes ASR generating annotations and ASR information stored in a header. James et al. does not describe header storage or language indications and Moshier does not describe language indications or header storage of word and phoneme sets identities. Bird et al., James et al. and Moshier do not provide evidence relevant to an objective teaching, motivation, or suggestion to construct such header information or to select and combine references holding the whole specific subject matter of this claims obvious to one of ordinary skill in the art of (data) corpus annotation, particularly including phoneme, phoneme link, word, word link, and a node storage area with both phoneme link and word link storage.

***Response to Arguments***

44. The prior Office action, mailed December 13, 2002 (paper 10), objects to the title, specification, and claims, and rejects claims under *35 USC § 112, § 101, § 102,* and *§ 103*. The Applicant's arguments and changes in AMENDMENT AND PETITION FOR EXTENSION OF TIME filed June 12, 2003 (paper 13) have been fully considered with the following results.

45. With respect to objection to the title, the new title is not sufficiently descriptive due to changes entered by amendment. Accordingly, the objection is maintained.

46. With respect to objection to the specification as failing to provide proper antecedent basis for data arranged in blocks of equal time duration, the Applicant's arguments appear to be as follows:

The Applicant's argument appears to be that the subject matter was cancelled from the claims. This argument is not persuasive because it appears in at least claim 99. The Applicant's

arguments have been fully considered, but they are not persuasive. Accordingly, the objection is maintained.

47. With respect to objection to the specification as failing to provide proper antecedent basis for the terminology "parol", the objection no longer applies because the subject matter has been canceled from the claims.

48. With respect to objections to the claims for requiring interpretation for claimed subject matter, the objections no longer apply because those claims have been canceled.

49. With respect to rejections of claims under *35 USC § 112, § 101, § 102,* and *§ 103*, the rejections no longer apply because the rejected claims have been canceled.

50. With respect to Bird et al., the Applicant's arguments appear to be as follows.

The Applicant's argument appears to be that claims distinguish from and are patentable over the reference by reciting storage areas containing phoneme and word data forming a temporal block of the annotation, with temporal in a time ordered sequence. This argument is not persuasive because the reference explicitly describes storage areas containing word data forming a temporal block of the annotation, storage areas containing phoneme data forming a temporal block of the annotation, and phoneme and word data forming a temporal block of the annotation, blocks having timing, storage area timing relative to the block. Bird et al. describes these features using the general file storage device and the list-of-data format that the specification [at Fig. 29 and pages 10-11] uses. In addition, Bird et al. [at sections 2.4, 4.8, and 8] invites the user to pick and

choose the relationships that are necessary, convenient, efficient, or enriching when the particular corpus and user's need make them obvious.

Please see new grounds of rejection.

***Conclusion***

51. The following reference here made of record is considered pertinent to applicant's disclosure:

Kaplan et al. [US Patent 5,594,641] claims a process of making a finite state transducer stored on a computer readable medium and a data structure stored on a computer readable medium that is processed by a program to search a database and retrieve information.

52. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

53. Any response to this action should be mailed to:

**Mail Stop AF**

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

**or faxed to:**

(703) 872-9315, (for formal communications; please mark "EXPEDITED PROCEDURE")

**Or:**

(703) 872-9315, (for informal or draft communications, and please label "PROPOSED" or "DRAFT")

Hand-delivered responses should be brought to Crystal Park II, 2121 Crystal Drive, Arlington, VA (Sixth Floor, Receptionist).

54. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Donald L. Storm, of Art Unit 2654, whose telephone number is (703)305-3941. The examiner can normally be reached on weekdays between 8:00 AM and 4:30 PM Eastern Time. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (703)305-9645. Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Technology Center 2600 Customer Service Office at telephone number (703)306-0377.

Donald L. Storm
Donald L. Storm
August 7, 2003

Richemond Dorvil
Primary Examiner